

# **1999 NEW YORK CITY HOUSING AND VACANCY SURVEY SAMPLE DESIGN, ESTIMATION PROCEDURE AND ACCURACY STATEMENT**

The purpose of the 1999 New York City Housing and Vacancy Survey (NYCHVS) is to measure rental and homeowner vacancy rates, as well as various household and person characteristics. New York City is required by law periodically to conduct such a survey in order to determine if rent regulations should be continued. The actual enumeration was between mid-January and mid-May of 1999.

## **I. SAMPLE DESIGN**

New York City's prime consideration is the "vacant available for rent" rate. This rate is the ratio of the vacant available for rent units to the total number of renter occupied and vacant available for rent units for the entire city. The design required the standard error of the estimate of this rate for the entire city be no more than one-fourth of 1 percent, if the actual rate was 3 percent.

We selected the sample from the following four sample frames:

1. Housing units included in the 1990 Census.
2. Housing units constructed since the 1990 Census.
3. Units which were nonresidential at the time of the 1990 Census, but have since been converted to housing units.
4. Housing units in structures owned by New York City (in rem) that were over-sampled in order for the City to learn more about the characteristics of occupants of these types of units. These housing units are all part of frame one with maybe a few in frame three.

### **A. Housing Units Included in the 1990 Census**

Within this frame, we sorted housing units by (a) borough, (b) sub-borough, (c) percent renter occupied in the block, (d) tract, (e) block number, (f) basic street address, and (g) unit designation. We selected a systematic sample of housing units across all boroughs. This frame included in rem units.

### **B. Housing Units Constructed Since the 1990 Census**

We selected units in this frame from Certificates of Occupancy (C of Os) issued between April 1, 1990 and October 31, 1998. We sorted the housing units on the C of Os by borough and date (i.e., year and month) of issue. We selected a systematic sample of housing units within each borough. We listed each structure containing sample housing units and assigned unit designations based on the order in which the sample housing unit appeared on these listings. We dropped all

sample housing units that were also on the 1990 census file from this sample.

#### C. Housing Units Converted from Non-residential Units Since the 1990 Census

Housing units in this frame were eligible for sample if a conversion C of O was issued for the structure between April 1, 1990 and October 31, 1998. We selected the sample from this frame using the same procedure as the frame for housing units constructed since the 1990 census.

#### D. Housing Units in Structures Owned by New York City (in rem)

This frame consisted of units in structures owned by New York City as of October 1998. The City owned these units because the owner failed to pay real estate tax and/or other charges on the property. In 1991, 1993, and 1996, we selected a supplemental sample of in rem units from the City's in rem lists at the time of each enumeration to supplement the sample of in rem units from the 1990 Census frame. In 1999, we kept units selected for the supplemental sample in previous enumerations if they were still on the City's in rem lists. Each enumeration, we dropped in rem units from prior years' supplemental samples which were no longer on the City's in rem lists. As a result, we needed to select additional supplemental in rem sample cases for 1999.

We selected the supplemental sample of in rem units in two steps. First, we sorted the buildings by:

- (1) enumeration the building was added to the City's in rem lists,
- (2) borough, and
- (3) size of the building (number of units)

and selected a systematic sample of buildings. Then, after listing the individual units in each building, we selected a systematic sample of units within each sample building.

#### E. Sample Size

Within each frame, we selected clusters (groups of housing units) of generally four housing units. For all frames except the in rem frame, the four housing units were consecutive units. For the in rem frame, we selected a systematic sample of the four housing units within each sample building.

The total number of sample housing units for New York City was 18,180. Of these housing units, 1,060 interviews were not obtained because, for occupied housing units,

- the occupants refused to be interviewed,
- were not at home after repeated visits,
- or were unavailable for some other reason.

For vacant units, an interview wasn't obtained if no informed respondent could be found after repeated visits. These 1,060 noninterviews are known as type-A noninterviews. This classification produced a 94-percent overall response rate. There were an additional 658 units, known as type-C noninterviews, that were not interviewed because they no longer exist or are uninhabitable. The table below provides the total number of sampled housing units by borough.

Borough	Number of Housing Units
Bronx	2,735
Brooklyn	5,146
Manhattan	4,916
Queens	4,504
Staten Island	879
Total	18,180

The sample housing units were visited in mid-January through mid-May 1999 by field representatives (FRs) hired and trained for this task. The FRs visited each sample address and completed a questionnaire for both occupied and vacant units. In addition, for evaluation purposes, the occupancy status of all vacant units and a sample of occupied units was independently determined in a reinterview. An independent third interview reconciled any differences.

#### F. Exclusions

The survey included only housing units. The principal exclusions were living quarters classified as:

- transient hotels,
- commercial and mission lodging houses,
- inmate living quarters in institutions,
- quarters for the military on military installations, and
- other large group quarters not meeting the definition of a housing unit.

Also generally excluded were housing units in special places. These included housing units located on the grounds of institutions (both civilian and military), with the exception of residential hotels and motels. Housing units with a Certificate of Occupancy issued for the first time after October 31, 1998 were also excluded.

## II. ESTIMATION PROCEDURE

To estimate housing unit characteristics we used a three-stage ratio estimation procedure. We used the same procedure to estimate person characteristics, but added a ratio estimate factor to adjust for person undercoverage within households. Prior to the ratio estimation procedures, we adjusted the basic weight (the inverse of the probability of selection for the unit) for each interviewed unit to account for type-A noninterviews.

In rem units had multiple chances of selection. They were eligible for selection from

1. both the 1990 Census and the respective in rem frames,
2. possibly the conversion frame (as such units could become in rem),
3. new construction frame (as such units could become in rem).

The basic weights reflect the fact that they had multiple chances of selection.

### A. Type-A Noninterview Adjustment Factor

We applied the noninterview adjustment factor (adjusting for type-A noninterviews) to all interviewed units separately for old construction units (frames one and four) and new construction/conversion units (frames two and three). For old construction units, we computed the factor separately by borough for 99 cells using the following NYCHVS characteristics:

(a) Monthly rent (less than 100, 100-199, 200-299, 300-399, 400-499, 500-599, 600-699, 700-999, 1000 +).

(b) Value (less than 25000, 25000-49999, 50000-74999, 75000-99999, 100000-149999, 150000-199999, 200000-249999, 250000-299999, 300000-399999, 400000-499999, 500000 +).

(c) Number of rooms (rent: 1, 2, 3, 4+, or 1-2, 3, 4, 5+, or 1-3, 4, 5, 6+; own: 1-4, 5, 6, 7+, or 1-3, 4, 5, 6+, or 1-3, 4, 5-6, 7+, or 1-4, 5, 6, 7+, or 1-5, 6, 7, 8+, or 1-5, 6-7, 8, 9+).

(d) Vacancy status (renter occupied/vacant for rent, owner occupied/vacant for sale, vacant/without tenure or vacancy status).

We used 1996 NYCHVS data, where available, to determine the tenure and characteristics of a unit. If the 1996 NYCHVS data were not available, we used either 1993 NYCHVS data or 1991 NYCHVS data or 1990 Census data or 1999 NYCHVS data (in that order). For new construction/conversion units, we computed the factor separately by type of unit (new construction and conversion), year the C of O was issued (new construction only), and borough. The noninterview adjustment factor was equal to the following ratio for each cell:

$$\frac{(\text{weighted count of interviewed units}) + (\text{weighted count of Type A noninterviews})}{(\text{weighted count of interviewed units})}$$

## B. Ratio Estimate Factors

We used a three-stage housing unit ratio estimation procedure and a one-stage person ratio estimation procedure:

- ! to account for known sampling variability in the 1990 census frame (frame one),
- ! to account for known sampling variability in the in rem frame (frame four),
- ! to bring the sample estimates of housing units into close agreement with estimates derived from independent sources, and
- ! to account for housing unit and person undercoverage.

For each ratio estimation procedure, we computed factors for ratio estimate cells (characteristics) and applied the factors to the appropriate units in the corresponding cell. The factors were equal to the following ratio:

$$\frac{\text{Independent estimate of number of HUs (persons) for the cell}}{\text{NYCHVS sample estimate of number of HUs (persons) for the cell}}$$

The denominators of the ratios equaled the sum of the weights of housing units, or persons, with all previous factors applied, on all records in the corresponding cell.

### 1. 1990 Census Ratio Estimate Factor

We applied this ratio estimation procedure to all 1990 census units in the NYCHVS sample (units from frame one and frame four). This procedure adjusted for differences between the 1990 census counts and the corresponding sample counts. These differences occurred because of sampling variability, which was increased since the sample was not selected from the final census file. We computed the factors separately by borough for each of the 138 cells using the following 1990 census characteristics: (a) monthly rent, (b) value, (c) race of householder, (d) Hispanic origin, (e) and vacancy status. We used the 1990 census counts of housing units as the independent estimates for each cell.

### 2. In Rem Ratio Estimate Factor

We applied this ratio estimation procedure to all in rem sample units (frames one and four). This procedure adjusted for known sampling variability in the in rem sample selection. We computed ratio estimate factors for each borough (5 cells). The independent estimates were the total number of in rem units in each borough in the in rem frame.

### 3. 1999 Housing Unit Ratio Estimate Factor

We applied this ratio estimation procedure to all interviewed housing units. This procedure adjusted the 1999 NYCHVS sample estimate for housing unit undercoverage by controlling the sample estimate to independently derived estimates. We calculated the ratio estimate factor for each of the boroughs (5 cells). The independent estimates were equal to the total number of housing units in each of the boroughs at the time of the survey.

### 4. 1999 Person Ratio Estimate Factor

This additional adjustment accounted for sampling variability and known coverage deficiencies for persons within interviewed households. We computed this factor within each borough by age, race and sex (80 cells).

During the analysis of the 1993 NYCHVS, inconsistencies were noted when estimates were made using different weights (i.e., housing unit weight vs. person or population weight) for certain person characteristics. In order to reduce the effect of this discrepancy, we modified the calculation of the person ratio estimate factor, beginning with the 1996 NYCHVS.

Previously, the ratio equaled the independent estimate of persons for the cell divided by the NYCHVS sample estimate of persons for the cell. This method assumes that all persons with a given age/race/sex have an equal chance of being missed by the survey. Some of the observed inconsistencies in the data could be eliminated by assuming that the reference person and his or her spouse or unmarried partner are always picked up by the survey if the housing unit is

interviewed (i.e., only persons other than reference persons, spouses, or unmarried partners could be missed in interviewed housing units).

Thus, the new numerator of the ratio equaled the independent estimate of persons for the cell minus the NYCHVS sample estimate of reference persons and spouses or unmarried partners. The new denominator of the ratio equaled the NYCHVS sample estimate of persons other than reference persons, spouses or unmarried partners for the cell. The new person ratio estimate factor was applied only to the persons other than reference persons, spouses, or unmarried partners.

### C. Change in Methodology to Compute Person Controls

For 1991, we extrapolated the change between the 1980 and 1990 censuses to derive the person controls. Beginning in 1993, we used independently derived current estimates based on the 1990 Census and Medicare data.

Since 1993, we computed controls using a modified 1990 Census age/race/sex classification. Among other things, the modified age/race/sex classification puts Hispanics whose race is classified as "other" into a specific race category. The 1993, 1996, and 1999 controls, based on a modified age/race/sex classification, reflect Hispanics in all race categories except "other". On the other hand, the 1991 controls reflect Hispanics in the "other race" category. Since this change caused unexplainable fluctuations in the "other race" category, we combined the white and "other race" categories in every enumeration since 1993 for the purposes of person ratio estimation.

As part of the regular NYCHVS processing, Hispanics and non-Hispanics who listed their race as "other" were allocated to specific race categories. Also, non-reports to the race question were allocated to specific race categories. The net effect of these changes was the African-American and "other race" controls increased and the white controls decreased. Some of this change may be real but most is probably due to the change in methodology.

The ratio estimation procedures, as well as the overall estimation procedure, reduced the sampling error for most statistics below what would have been obtained by simply weighting the sample by the basic weight.

## **III. SAMPLING AND NONSAMPLING ERRORS**

The statistics produced from this survey are estimates derived from a sample. They will differ from the true values being estimated. There are two types of errors which cause estimates based on a sample survey to differ from the true value - sampling error and nonsampling error.

### A. Nonsampling Errors

Suppose every housing unit in New York City were interviewed. Estimates would still differ from the true value (for example, the median contract rent). In this instance, the difference is due solely to nonsampling errors. We attribute nonsampling errors in sample surveys to many sources:

- ! deficiencies in the sampling frame (i.e., not all housing units are covered),
- ! inability to pick up all persons within sample households,
- ! inability to obtain information about all cases in the sample,
- ! definitional difficulties,
- ! differences in the interpretation of questions,
- ! inability or unwillingness to provide correct information on the part of the respondents, and
- ! mistakes in recording, coding or keying the data obtained.

There are also other errors of collection, response, processing, coverage, and estimation for missing data.

In the 1999 NYCHVS, we missed about three-fourths of 1 percent of the housing units in the five boroughs covered by the survey. Overall, we missed about 3 percent of the people in sample households. This within-household undercoverage varied by age, race, sex, and borough. It ranged from about a 37-percent overcoverage of African American females between 15-24 in Staten Island to a 32-percent undercoverage of African American males between 25-64 in Manhattan. The following table gives the undercoverage of the various race-sex groups for the city as a whole:

Race-Sex Group	Undercoverage
White & Other Females	.014 %
White & Other Males	2 %
African American Females	5 %
African American Males	9 %

We adjusted for this undercoverage through the housing unit and person ratio estimate factors. Measures of other errors for this survey are not available. However, we believe some of the important response and most of the operational errors were detected and corrected during the Bureau's review of the data for reasonableness and consistency.

#### B. Sampling Errors

Sampling error reflects how estimates from a sample vary from the actual value. **NOTE: By the term "actual value" we mean the value we would have gotten had all housing units been interviewed, under the same conditions, rather than only a sample.**

The formulas in Tables 1 through 6 allow you to compute a range of error such that there is a known probability of being correct if you say the actual value is within the range. The error formulas are approximations to the errors. They indicate the order of magnitude of the errors rather than the actual errors for any specific characteristic. To construct the range, add and subtract the error computed from the formulas to the estimate.

The letter "A" in the formula represents the weighted sample estimate you derive from the file.

The letter "Z" determines the probability the actual value is within the range you compute. The larger the value of Z, the larger the range, and the higher the odds the actual value will be in the range. The following values of Z are most commonly used.

Value of Z	Meaning
1.00	There is a 67-percent chance you'll be correct if you say the actual value is in the range you compute.
1.64	There is a 90-percent chance you'll be correct if you say the actual value is in the range you compute.
1.96	There is a 95-percent chance you'll be correct if you say the actual value is in the range you compute.
2.58	There is a 99-percent chance you'll be correct if you say the actual value is in the range you compute.

Note that if  $Z = 1.00$ , the formula computes the standard error. Ranges of 90 and 95-percent are commonly used. The range of error is also referred to as the confidence interval since there is a certain level of confidence the actual value is within the interval.

For example there are 19,819 vacant-for-rent units in Brooklyn. To compute a 90-percent confidence interval, you would use the first formula in Table 3 and you would compute the error

$$Z \times \sqrt{(266.27 \times A) - (.000307 \times A^2)}$$

$$1.64 \times \sqrt{(266.27 \times 19,819) - (.000307 \times 19,819^2)} = 3,724$$

as follows:

Thus there is a 90-percent chance you'll be correct if you conclude the actual number of vacant-for-rent units in Brooklyn is 19,819 plus or minus 3,724, or in the range 16,095 to 23,543.

If the estimate involves two characteristics from Tables 1 through 6, use the formula with the larger first number under the square root.

## 1. Percents

$$Z \times Y \times \sqrt{\frac{266.27 \times P \times (100 - P)}{B}}$$

The formula for computing the error of any percent derived from the data is the following:  
where:

Z: defines the confidence the range will include the actual value,

Y: is the number from the last column of Tables 1 through 6 (chosen based on the denominator),

P: is the percent you calculate, and

B: is the denominator of the percent.

For example, there are 898,395 households in units built between 1947 and 1969 and 338,701, or 37.7 percent, are owners. To compute a 90-percent confidence interval you would plug the following numbers into the above formula:

$$1.64 \times 1.189 \times \sqrt{\frac{266.27 \times 37.7 \times 62.3}{898,395}} = 1.6$$

Thus, if you say that the actual percentage of owners in buildings built between 1947 and 1969 is between 36.1 percent and 39.3 percent, there is a 90-percent chance you'll be correct.

## 2. Differences

People often ask whether two numbers are actually different. If the range of error for the difference doesn't include zero, the numbers are different. As a general rule, if the confidence intervals don't overlap, they're different. To compute the range of error of the difference use the

$$\sqrt{(\text{error on first number})^2 + (\text{error on second number})^2}$$

following formula:

This formula is quite accurate for (a) the difference between estimates of the same item in two different areas or (b) the difference between separate and uncorrelated items in the same area. If there is a high positive correlation between the two items, the formula will overestimate the error.

If there is a high negative correlation, the formula will underestimate the error. The following illustration shows how to compute the error of a difference.

There are 10,406 vacant-for-rent units in New York City with 3 to 5 units in the building and 3,646 vacant-for-rent units with 6 to 9 units in the building. The respective errors for a 90-percent confidence interval are 2,725 and 1,615. The error for a 90-percent confidence interval for the 6,760 difference is the following:

$$\sqrt{(2,725)^2 + (1,615)^2} = 3,168$$

Thus, there is a 90-percent chance you'll be correct if you say the actual difference between vacant-for-rent units in 3 to 5 unit buildings vs. 6 to 9 unit buildings in New York City is between 3,592 and 9,928.

### 3. Medians

The median is the value 50-percent of the way through the distribution. Thus, 50-percent of the total falls below and 50-percent falls above the median. Note that the median presented in this example is the true median (i.e., computed by SAS) not an approximation. You can construct a confidence interval around the median by computing the standard error on a 50-percent characteristic and then translating that into an interval for the characteristic.

- a. Using the error formula for percents, above, compute the error of 50-percent. The total number of housing units from the distribution is the denominator in the formula. Subtract the "not applicable" category from the total.
- b. Calculate the confidence interval for the true median by adding and subtracting the width of the interval containing the median times the standard error on the 50-percent characteristic divided by the proportion of units in the interval containing the median, to the median.

The probability you will be correct if you conclude that the actual median is within the interval depends on the value of Z in the error of percent formula. The following example shows how to compute a 90-percent confidence interval.

The median value for all occupied housing units in New York City is \$190,000. The number of occupied housing units in the distribution of value of units is presented below.

Distribution of Value of Units

Value	Number of HUs
-------	---------------

Less Than \$25,000	61,739
\$25,000-\$49,999	40,434
\$50,000-\$74,999	43,832
\$75,000-\$99,999	31,103
\$100,000-\$149,999	93,692
\$150,000-\$199,999	203,682
\$200,000-\$249,999	154,846
\$250,000-\$299,999	108,783
\$300,000-\$349,999	56,393
\$350,000-\$399,999	28,257
\$400,000-\$499,999	29,249
\$500,000-\$599,999	14,890
\$600,000-\$699,999	6,396
\$700,000-\$799,999	11,246
\$800,000-\$999,999	10,112
\$1,000,000 or more	20,474
Not Applicable	1,953,289
TOTAL	2,868,412

1. The error on a 50-percent characteristic based on 915,123 (2,868,412 minus the "not applicable" number) housing units is calculated as follows:

$$1.64 \times 1.0000 \times \sqrt{\frac{266.27 \times 50 \times 50}{915,123}} = 1.40$$

2. The 90-percent confidence interval for the median (\$190,000) is:

$$190,000 \pm (199,999.5 - 149,999.5) \times \frac{1.40}{22.26} = 190,000 \pm 3,145$$

where:

- ! 199,999.5-149,999.5 is the width of the interval that contains the median
- ! 1.40 is the error for a 90-percent confidence interval for the 50-percent characteristic
- ! 22.26 is the percent of cases that fall in the interval containing the median

Thus, there is a 90-percent chance that you will be correct if you conclude that the actual median for all occupied housing units in New York City (\$190,000) is between \$186,855 and \$193,145.

#### 4. Means

The mean and the median usually differ. The mean is usually higher because it is influenced more heavily than the median by very large values. Use the following formula to estimate the error of the mean:

$$Z \times Y \times \sqrt{\frac{\sum_{i=1}^n p_i x_i^2 - (\sum_{i=1}^n p_i x_i)^2}{c}} \times 266.27$$

where:

Y: is the number from the last column of Tables 1 through 6

Z: defines the confidence the range will include the actual value

$p_i$ : is the proportion of total households or persons from a distribution in the  $i^{\text{th}}$  interval

$x_i$ : is the midpoint of the  $i^{\text{th}}$  interval (NOTE: The midpoint of the open-ended interval is 1.5 times the lower limit)

$c$ : is the total number of households or persons in the distribution (NOTE: Subtract the number of "not applicable" from the total to get  $c$ )

$n$ : is the total number of intervals in the distribution

For example, the mean (or average) value of all owner-occupied housing units in New York City was \$235,358 (compared to a median of \$190,000). The distribution from which the mean was computed is given below.

Value	Number of HUs	$p_i$	$x_i$
Less Than \$25,000	61,739	.0675	\$12,500
\$25,000-\$49,999	40,434	.0442	\$37,500
\$50,000-\$74,999	43,832	.0479	\$62,500
\$75,000-\$99,999	31,103	.0340	\$87,500
\$100,000-\$149,999	93,692	.1024	\$125,000
\$150,000-\$199,999	203,682	.2226	\$175,000
\$200,000-\$249,999	154,846	.1692	\$225,000
\$250,000-\$299,999	108,783	.1189	\$275,000
\$300,000-\$349,999	56,393	.0616	\$325,000
\$350,000-\$399,999	28,257	.0309	\$375,000
\$400,000-\$499,999	29,249	.0320	\$450,000
\$500,000-\$599,999	14,890	.0163	\$550,000
\$600,000-\$699,999	6,396	.0070	\$650,000
\$700,000-\$799,999	11,246	.0123	\$750,000
\$800,000-\$999,999	10,112	.0110	\$900,000
\$1,000,000 Or More	20,474	.0224	\$1,500,000
Not Applicable	1,953,289	----	
Total	2,868,412	1.000	

Plugging the numbers in the above formula, the error for a 90-percent confidence interval on the mean income is computed as follows:

$$1.64 \times 1.000 \times \sqrt{\frac{117,943,625,000 - (243,240.00)^2}{915,123}} \times 266.27 = \$6,782$$

Thus, there is a 90-percent chance of being correct if you say the mean value of all owner-occupied housing units in New York City was between \$228,576 and \$242,140.

**Table 1: Errors for New York City**

	Publication Estimates	Percentages
	The error is the larger of:	Value of Y for Percent Formula
<b>Errors on Housing Units</b>		
Characteristics of Housing Units Not Listed Below	$Z \times \sqrt{266.27 \times A - .000088 \times A^2}$ or $Z \times 266$	1.000
Housing Unit Clustering Items <sup>1</sup> (see Table 7 for a listing)	$Z \times \sqrt{360.62 \times A - .000118 \times A^2}$ or $Z \times 361$	1.164
NYC Housing Unit totals (all borough and sub-borough)	$Z \times \sqrt{534.71 \times A - .000165 \times A^2}$ or $Z \times 535$	1.417
<b>Errors on Persons</b>		
Characteristics of Persons Not Listed Below	$Z \times \sqrt{304.11 \times A - .000042 \times A^2}$ or $Z \times 304$	1.069
	<b>NOTE: For any of the person characteristics listed below that are cross-tabbed by Borough and Sub-borough use the formula for the specific characteristic listed below. Don't use the formulas listed below for cross-tabs of characteristics of persons listed below {e.g., Age by sex (males under 25), Age by Race (African Americans under 25), or sex by race (white females)}. Use the formula above (Characteristics of Persons Not Listed Below).</b>	
<b>Whites and other Races and Ethnicity</b>	$Z \times \sqrt{813.49 \times A - .000162 \times A^2}$ or $Z \times 813$	1.748
<b>Males</b>	$Z \times \sqrt{813.49 \times A - .000239 \times A^2}$ or $Z \times 813$	1.748
<b>Females</b>	$Z \times \sqrt{813.49 \times A - .000212 \times A^2}$ or $Z \times 813$	1.748
<b>Persons under 25 yrs. old</b>	$Z \times \sqrt{531.86 \times A - .000073 \times A^2}$ or $Z \times 532$	1.413
<b>African Americans</b>	$Z \times \sqrt{1,533.89 \times A - .000687 \times A^2}$ or $Z \times 1,534$	2.400
<b>Borough and Sub-borough</b>	$Z \times \sqrt{1,533.89 \times A - .000212 \times A^2}$ or $Z \times 1,534$	2.400

<sup>1</sup> Clustering Items are items that tend to apply to all units in a building.

**Table 2: Errors for Bronx**

	Publication Estimates	Percentages
	The error is the larger of:	Value of Y for Percent Formula
<b>Errors on Housing Units</b>		
<b>Characteristics of Housing Units Not Listed Below</b>	$Z \times \sqrt{266.27 \times A - .000593 \times A^2}$ or $Z \times 266$	1.000
<b>Housing Unit Clustering Items<sup>1</sup> (see Table 7 for a listing)</b>	$Z \times \sqrt{360.62 \times A - .000799 \times A^2}$ or $Z \times 361$	1.164
Sub-borough and Borough Housing Unit totals	$Z \times \sqrt{534.71 \times A - .001190 \times A^2}$ or $Z \times 534$	1.417
<b>Errors on Persons</b>		
<b>Characteristics of Persons Not Listed Below</b>	$Z \times \sqrt{304.11 \times A - .000268 \times A^2}$ or $Z \times 304$	1.069
	<b>NOTE: For any of the person characteristics listed below that are cross-tabbed by Borough and Sub-borough use the formula for the specific characteristic listed below. Don't use the formulas listed below for cross-tabs of characteristics of persons listed below {e.g., Age by sex (males under 25), Age by Race (African Americans under 25), or sex by race (white females)}. Use the formula above (Characteristics of Persons Not Listed Below).</b>	
<b>Whites and other Races and Ethnicity</b>	$Z \times \sqrt{813.49 \times A - .001252 \times A^2}$ or $Z \times 813$	1.748
<b>Males</b>	$Z \times \sqrt{813.49 \times A - .001568 \times A^2}$ or $Z \times 813$	1.748
<b>Females</b>	$Z \times \sqrt{813.49 \times A - .001325 \times A^2}$ or $Z \times 813$	1.748
<b>Persons under 25 yrs. old</b>	$Z \times \sqrt{531.86 \times A - .000469 \times A^2}$ or $Z \times 532$	1.413
<b>African Americans</b>	$Z \times \sqrt{1,533.89 \times A - .003174 \times A^2}$ or $Z \times 1,534$	2.400
<b>Sub-borough and Borough</b>	$Z \times \sqrt{1,533.89 \times A - .001354 \times A^2}$ or $Z \times 1,534$	2.400

<sup>1</sup> Clustering Items are items that tend to apply to all units in a building.

**Table 3: Errors for Brooklyn**

	Publication Estimates	Percentages
	The error is the larger of:	Value of Y for Percent Formula
<b>Errors on Housing Units</b>		
Characteristics of Housing Units Not Listed Below	$Z \times \sqrt{266.27 \times A - .000307 \times A^2}$ or $Z \times 266$	1.000
<b>Housing Unit Clustering Items<sup>1</sup> (see Table 7 for a listing)</b>	$Z \times \sqrt{360.62 \times A - .000413 \times A^2}$ or $Z \times 361$	1.164
Sub-borough and Borough Housing Unit totals	$Z \times \sqrt{534.71 \times A - .000616 \times A^2}$ or $Z \times 535$	1.417
<b>Errors on Persons</b>		
<b>Characteristics of Persons Not Listed Below</b>	$Z \times \sqrt{304.11 \times A - .000138 \times A^2}$ or $Z \times 295$	1.069
	<b>NOTE: For any of the person characteristics listed below that are cross-tabbed by Borough and Sub-borough use the formula for the specific characteristic listed below. Don't use the formulas listed below for cross-tabs of characteristics of persons listed below {e.g., Age by sex (males under 25), Age by Race (African Americans under 25), or sex by race (white females)}. Use the formula above (Characteristics of Persons Not Listed Below).</b>	
<b>Whites and other Races and Ethnicity</b>	$Z \times \sqrt{813.49 \times A - .000615 \times A^2}$ or $Z \times 813$	1.748
<b>Males</b>	$Z \times \sqrt{813.49 \times A - .000789 \times A^2}$ or $Z \times 813$	1.748
<b>Females</b>	$Z \times \sqrt{813.49 \times A - .000691 \times A^2}$ or $Z \times 813$	1.748
<b>Persons under 25 yrs. old</b>	$Z \times \sqrt{531.86 \times A - .000241 \times A^2}$ or $Z \times 532$	1.413
<b>African Americans</b>	$Z \times \sqrt{1,533.89 \times A - .001731 \times A^2}$ or $Z \times 1,534$	2.400
<b>Sub-borough and Borough</b>	$Z \times \sqrt{1,533.89 \times A - .000694 \times A^2}$ or $Z \times 1,534$	2.400

<sup>1</sup> Clustering Items are items that tend to apply to all units in a building.

**Table 4: Errors for Manhattan**

	Publication Estimates	Percentages
	The error is the larger of:	Value of Y for Percent Formula
<b>Errors on Housing Units</b>		
<b>Characteristics of Housing Units Not Listed Below</b>	$Z \times \sqrt{266.27 \times A - .000340 \times A^2}$ or $Z \times 266$	1.0000
<b>Housing Unit Clustering Items<sup>1</sup></b> (see Table 7 for a listing)	$Z \times \sqrt{360.62 \times A - .000459 \times A^2}$ or $Z \times 361$	1.164
Sub-borough and Borough Housing Unit totals	$Z \times \sqrt{534.71 \times A - .000684 \times A^2}$ or $Z \times 535$	1.417
<b>Errors on Persons</b>		
<b>Characteristics of Persons Not Listed Below</b>	$Z \times \sqrt{304.11 \times A - .000196 \times A^2}$ or $Z \times 295$	1.069
	<b>NOTE: For any of the person characteristics listed below that are cross-tabbed by Borough and Sub-borough use the formula for the specific characteristic listed below. Don't use the formulas listed below for cross-tabs of characteristics of persons listed below {e.g., Age by sex (males under 25), Age by Race (African Americans under 25), or sex by race (white females)}. Use the formula above (Characteristics of Persons Not Listed Below).</b>	
<b>Whites and other Races and Ethnicity</b>	$Z \times \sqrt{813.49 \times A - .000698 \times A^2}$ or $Z \times 813$	1.748
<b>Males</b>	$Z \times \sqrt{813.49 \times A - .001112 \times A^2}$ or $Z \times 813$	1.748
<b>Females</b>	$Z \times \sqrt{813.49 \times A - .000991 \times A^2}$ or $Z \times 813$	1.748
<b>Persons under 25 yrs. old</b>	$Z \times \sqrt{531.86 \times A - .000343 \times A^2}$ or $Z \times 532$	1.413
<b>African Americans</b>	$Z \times \sqrt{1,533.89 \times A - .003972 \times A^2}$ or $Z \times 1,534$	2.400
<b>Sub-borough and Borough</b>	$Z \times \sqrt{1,533.89 \times A - .000988 \times A^2}$ or $Z \times 1,534$	2.400

<sup>1</sup> Clustering Items are items that tend to apply to all units in a building.

**Table 5: Errors for Queens**

	Publication Estimates	Percentages
	The error is the larger of:	Value of Y for Percent Formula
<b>Errors on Housing Units</b>		
<b>Characteristics of Housing Units Not Listed Below</b>	$Z \times \sqrt{266.27 \times A - .000339 \times A^2}$ or $Z \times 266$	1.000
<b>Housing Unit Clustering Items<sup>1</sup> (see Table 7 for a listing)</b>	$Z \times \sqrt{360.62 \times A - .000456 \times A^2}$ or $Z \times 361$	1.164
Sub-borough and Borough Housing Unit totals	$Z \times \sqrt{534.71 \times A - .000680 \times A^2}$ or $Z \times 535$	1.417
<b>Errors on Persons</b>		
<b>Characteristics of Persons Not Listed Below</b>	$Z \times \sqrt{304.11 \times A - .000156 \times A^2}$ or $Z \times 304$	1.069
	<b>NOTE: For any of the person characteristics listed below that are cross-tabbed by Borough and Sub-borough use the formula for the specific characteristic listed below. Don't use the formulas listed below for cross-tabs of characteristics of persons listed below {e.g., Age by sex (males under 25), Age by Race (African Americans under 25), or sex by race (white females)}. Use the formula above (Characteristics of Persons Not Listed Below).</b>	
<b>Whites and other Races and Ethnicity</b>	$Z \times \sqrt{813.49 \times A - .000539 \times A^2}$ or $Z \times 813$	1.748
<b>Males</b>	$Z \times \sqrt{813.49 \times A - .000871 \times A^2}$ or $Z \times 813$	1.748
<b>Females</b>	$Z \times \sqrt{813.49 \times A - .000799 \times A^2}$ or $Z \times 813$	1.748
<b>Persons under 25 yrs. old</b>	$Z \times \sqrt{531.86 \times A - .000272 \times A^2}$ or $Z \times 532$	1.413
<b>African Americans</b>	$Z \times \sqrt{1,533.89 \times A - .003458 \times A^2}$ or $Z \times 1,534$	2.400
<b>Sub-borough and Borough</b>	$Z \times \sqrt{1,533.89 \times A - .000786 \times A^2}$ or $Z \times 1,534$	2.400

<sup>1</sup> Clustering Items are items that tend to apply to all units in a building.

**Table 6: Errors for Staten Island**

	Publication Estimates	Percentages
	The error is the larger of:	Value of Y for Percent Formula
<b>Errors on Housing Units</b>		
<b>Characteristics of Housing Units Not Listed Below</b>	$Z \times \sqrt{266.27 \times A - .001744 \times A^2}$ or $Z \times 266$	1.000
<b>Housing Unit Clustering Items<sup>1</sup> (see Table 7 for a listing)</b>	$Z \times \sqrt{360.62 \times A - .002350 \times A^2}$ or $Z \times 361$	1.164
Sub-borough and Borough Housing Unit totals	$Z \times \sqrt{534.71 \times A - .003502 \times A^2}$ or $Z \times 535$	1.417
<b>Errors on Persons</b>		
<b>Characteristics of Persons Not Listed Below</b>	$Z \times \sqrt{304.11 \times A - .000763 \times A^2}$ or $Z \times 304$	1.069
	<b>NOTE: For any of the person characteristics listed below that are cross-tabbed by Borough and Sub-borough use the formula for the specific characteristic listed below. Don't use the formulas listed below for cross-tabs of characteristics of persons listed below {e.g., Age by sex (males under 25), Age by Race (African Americans under 25), or sex by race (white females)}. Use the formula above (Characteristics of Persons Not Listed Below).</b>	
<b>Whites and other Races and Ethnicity</b>	$Z \times \sqrt{813.49 \times A - .002219 \times A^2}$ or $Z \times 813$	1.803
<b>Males</b>	$Z \times \sqrt{813.49 \times A - .004212 \times A^2}$ or $Z \times 813$	1.803
<b>Females</b>	$Z \times \sqrt{813.49 \times A - .003958 \times A^2}$ or $Z \times 813$	1.803
<b>Persons under 25 yrs. old</b>	$Z \times \sqrt{531.86 \times A - .001334 \times A^2}$ or $Z \times 532$	1.413
<b>African Americans</b>	$Z \times \sqrt{1,533.89 \times A - .047835 \times A^2}$ or $Z \times 1,534$	2.400
<b>Sub-borough and Borough</b>	$Z \times \sqrt{1,533.89 \times A - .003848 \times A^2}$ or $Z \times 1,534$	2.400

<sup>1</sup> Clustering Items are items that tend to apply to all units in a building.

Table 7

### **Housing Unit Clustering Items**

- ! Access from Sidewalk to Elevator/Unit without using Stairs
- ! Additional Heating Required
- ! Boarded up Buildings in Neighborhood
- ! Broken Plaster/Peeling Paint
- ! Condition of Building and External Walls, Windows, Stairways, and Floors of Building for Total Occupied and Renter Occupied
- ! Control Status (renters and owners)
- ! Elevator in Building with 2 or more stories
- ! Floor Unit is on
- ! Heating Fuel - utility gas electricity only
- ! Heating System Breakdown
- ! Households Receiving Public Assistance/Welfare Payments
- ! Length of Lease
- ! Maintenance Deficiencies
- ! Number of Stories in Building.
- ! Number of Units in Building
- ! Plumbing Facilities
- ! Poor physical condition of Building
- ! Race and Ethnicity of Householder
- ! Rodent Infestation
- ! Structure Classification/Condition Rating
- ! Wheel Chair Accessibility
- ! Year Building Built